

シンポジウム

癌の機能ゲノミクス解析

油谷 浩幸

(東京大学先端科学技術研究センター)



座長 久武幸司(埼玉医科大学分子生物学)

分子生物学の久武です。次の演題は油谷浩幸先生で、「癌の機能ゲノミクス解析」という演題でお話をしてもらいます。油谷先生の経歴を簡単にご紹介しますと、先生は1980年に東京大学医学部を卒業されています。そのあと研修をされ、第3内科に入られました。第3内科から1988年にアメリカのMITの癌研究センターに留学されています。1994年に日本に帰られ、そのあと東京大学にある先端科学技術研究センターに最初は助教授として、今年からは教授として仕事をされています。今の研究テーマとして、癌のシステム生物学を研究されています。

DNAチップを用いて遺伝子発現を調べる方法が、最近によくやられています。我々のように昔から分子生物学をやっている人は、遺伝子の発現というとノーザンブロット法をやっていましたが、最近はDNAチップを使って数百、数千、さらには万の単位で遺伝子の発現を調べることが行われています。今日の先生のお話は、そういう技術を使って癌の性質を調べるということです。最新の非常におもしろいお話が聞けるのではないかと思います。



今日はゲノム医学研究センターの開設にあたり、こういう会でお話をさせていただき機会をいただきましてありがとうございます。村松先生には研修のころからいろいろと研究のご相談などに乗っていただき、なつかしい思いがします。

すでに林崎先生がマイクロアレイのお話をされていますし、私どもはシステムは違いますが、マイクロアレイとはどういうものかも含めて若干ご紹介申し上げたいと思います。

先程の村松先生のお話にもありましたように、生命のセントラルドグマというものがあり、DNAのインフォメーションがRNAに伝わって、RNAからタンパク質になるという中で、最近の学問はとにかく網羅的に見ようと。つまり、患者を診るときも、漢方ではありませんが、全体を見ることが非常に大事になってきます。これまでのように遺伝子一つ一つの形を

見る、機能を見るということではなくて、その総体を見ることが大事になってきます。その学問はオーミックス、いわゆるゲノム、トランスクリプトーム、プロテオームという言葉がだんだんこの科学の場で氾濫してきます。

オーミックスという場合に、ゲノムがDNAの総体であれば、そのRNAの転写されたプロファイルの全体像がトランスクリプトームであり、プロテインの総体がプロテオームです。プロテオーム解析については、今後、技術的なブレイクスルーがないと、まだまだ何十万というタンパク質の全貌を得ることは難しいのです。そういう意味で、現在ある数万の遺伝子の全体の少なくとも発現の量を見ることは、十分に可能になってきたということです。

そういうことをすると何ができるかですが、その遺伝子がどここの組織に出ているかで、病気の起きている細胞で、遺伝子の発現のプロファイルを調べることで、病態の理解につながることを期待されます。また、癌細胞などでの変化を見れば、例えばAPC (Adenomatous Polyposis Coli) という大腸癌の遺伝子

は、東大医科研の中村先生などがVogelsteinらと一緒にとられたわけですが、こういうものは癌抑制遺伝子ですから、発現が低下していることが認められます。

あるいはMycというoncogeneであれば、これは8q24にある遺伝子ですが、発現が非常に亢進しているということが見られます。この場合にはDNAのレベルでも普通、増幅していることが多いわけですが、発現量が非常に亢進しています。rasのようにポイントミューテーションでactivation（活性化）が起きるような遺伝子の場合には、発現レベルにはそれほど異常はないけれども、活性化が認められます。癌という病態はこういうさまざまな遺伝子変化の結果として出来上がってくるわけです。

その全体像を見るDNAチップとは何かということですが、定義としては遺伝子のクローン、cDNAのクローンを非常に高精度に高密度に配列化したものです。最近ではDNA合成機の精度も上がり、100ベースぐらいのものであればin vitroで作れることになっていますから、そういう合成のDNAを配列することも可能です。塩基配列を調べる、遺伝子の発現量を調べる、または多型を調べる、染色体の数を調べるといったさまざまな応用が可能です。今日は癌の分類という話を中心に申し上げますが、将来的にこれが治療法の選択にもつながります。多型の解析に使われるだろうということが期待されますし、臨床検査の分野ではO-157など病原性の大腸菌なのかどうかという検査にも応用が十分に期待されます。

現在、私たちが使っているプラットフォームとしては、ジーンチップのシステムと、マイクロアレイという、スライド上にクローンをスポットしていくタイプのテクノロジーの、両方があります。それぞれ長所と短所があり、ヒトの遺伝子であれば、現在は数万の遺伝子が解析できます。これは既知の遺伝子の数プラスEST(Expressed Sequence Tag)ということで、実際の遺伝子のユニットとしてのローカスの数よりは多くなっています。マウスにおいても、現在のシステムを使えば3万6000ということになります。

これはマイクロアレイの実験の行程についてどこかのウェブサイトから取ってきたものですが、ガラス上にこういうピン、あるいは現在はインクジェットの機械を用いてスポットングを行っていきます。そして物理的に、普通に実験に使うスライドガラスに、1枚2万というスポットで遺伝子を貼り付けていくわけです。そのときに対照サンプルと、自分が調べたいサンプルを別々の蛍光色素で、例えば赤と緑に標識すれば、それによって、あとはどちらのサンプルで遺伝子

が増えているかがわかります。

技術的に機材としてはこういうスポットターを用います。拡大するとこのような小さいピンが並んでおり、このピンで遺伝子を打っていくわけです。この中にガラスのスライドをたくさん並べておけば、一度に何十枚ものスライドが作れ、そのあとはこのようなhybridizationの装置を用います。この機械はうちの研究室にあるもので、12枚が使える、この中でハイブリダイゼーションと洗浄を行います。そのあとは共焦点のレーザースキャナーでスキャンを行います。これはハイブリダイゼーション装置の拡大です。これはスキャンしたイメージで、通常はCy3, Cy5という励起波長の異なる蛍光色素が使われます。

一方、アフィメトリックス社のシステムであるジーンチップはほとんど既製品ですが、半導体加工技術である光リソグラフィという技術で、光マスク（フォトマスク）の穴のあいているところだけに光が当たります。その光が当たったところだけ化学反応が起きることを用いて、シリコンの基板上に約20~25ミクロン四方の小さい単位で、別々の異なったDNAを合成します。そうすると、現在は40万種類ぐらいのDNAがここで合成でき、大体1枚のアレイで、1万~1万2000個の遺伝子の発現プロファイルが測定できます。

これがスキャンしたイメージで、先程のスポットとは違い、一つ一つの光マスクの形が正方形になっています。このように全く発現していない場合にはほとんどシグナルがないけれども、発現が認められる、ハイブリダイゼーションが行われた場合にはシグナルが得られます。2段になっているのは、似たような配列を下に配置しておくことによって、ミスマッチを用いた配列に対して、それは一応バックグラウンドで差し引けるようなシステムになっています。

遺伝子のなるべく3'側の、ほかの遺伝子とホモロジーのない部分に、このように25ベースのプロープ配列をデザインして、1つの遺伝子を16ぐらいのプロープペアで代表させて測定を行います。

今日はこのようなシステムを用いての腫瘍の診断、あるいは治療法の選択ということで、まずは肝細胞癌の例を説明します。用いた腫瘍は肝細胞癌、小児腫瘍、この対照になる非癌部および正常肝です。これは大腸癌患者の肝転移巣の非癌部です。つまり、こちらは肝炎ウイルスの感染を伴った肝腫瘍の非癌部組織です。それをアフィメトリックス社から現在売られているU95アレイで解析するわけです。これはUnigeneのビルド95というバージョンに合わせて作ったもので、1万2000個のヒトの遺伝子が解析できます。

癌と癌でない部分の区別はそれほど難しいもので

はありません。こちら側の上半分は非癌部で、下側が癌部ですが、このように明らかに発現のプロファイル、一つ一つの列のバーコードのパターンが遺伝子が上がっている下がっていることを示します。そしてこの一つ一つが遺伝子で、ここには1万2000個のうち変化のあった3000個ぐらいの遺伝子について、示してあります。このパターンによって仕分けると、癌の方では発現が増えている遺伝子もありますし、癌になることによって発現が低下する遺伝子もありますから、このように明らかにパターンが違うことがわかります。

その内訳を見ていきますと、こちらの正常、大腸癌肝転移、あるいは肝炎患者の非癌部組織もほぼ分かれていますし、小児腫瘍はここに分かれていますし、この上が肝細胞癌です。こういう単純な二次元の2方向のクラスタリングにおいてもこの程度の分類は可能ですが、さらにこの中に未分化型の肝癌、および中分化型、高分化型という肝癌の細かい分類を試みようとしますと、未分化型の肝癌は小児の肝芽腫の未分化型 (embryonal type) と非常に近いということもあり、このクラスタリングのパターンが非常に複雑になってきます。

ここでは、別の多変量解析の1つの手法である主成分分析法を用いて、三次元上に第3成分までを表示して、パターンとして分類できるかを表示してあります。外側のこのあたりに見えますのが小児の腫瘍であり、この赤いところが高分化型、中分化型の肝癌、そしてこちら側に正常群があります。そういう目で見れば、大体グループを形成しているのがわかります。

ここでは正常肝、慢性肝炎と、こちら側が非癌部組織、こちら側が癌部です。約1~2例のはぐれものだけが逆になっていますが、分類は十分に可能です。ただ、この中で慢性肝炎患者と正常肝は、先程のクラスター解析と同様に区別が可能です。

肝細胞癌の場合には日本の人口の約2%がHCV感染陽性ですし、1%強がHBV感染陽性ですが、当然、その両方から肝癌が発症してくるわけです。その違いによる差がプロファイルにあるかを見てみますと、癌部に関してはこの水色のHCVの患者と黄色のHBVの方ではあまり差がありません。一方、こちら側では、ある程度水色の群と黄色の群がまとまっている傾向が認められます。このことについてはこのあとにまた戻ってきますので、この差はいったい何だろうかということについてご説明を申し上げます。

先程の林崎先生のお話にもありましたが、ゲノム情報、ゲノムの配列はすでにかなり決まっています。ですから、こういう発現情報とゲノム情報の間

に何か関連を見いだせないかということで、私どももゲノム情報上にこういう発現情報を関係づけることを試んでいます。

癌がどんどん進行していくうえで、染色体が一部欠けていくことが認められます。あるいは、先程のMycという遺伝子のように増幅が起こることがあります。これはDNAの上で確実に変化が起きているわけですが、一方、epigenetics, imprintingという状態で代表される、DNA配列自体には何も変化が起きていないけれども、遺伝子の発現レベルが増えたり減ったりするということが実際に起こっています。癌ではこのように状態が狂っている、暴走してしまっていることがしばしば認められます。ということで、これは従来の例えば染色体が欠けているかという解析では見いだせないわけです。そこで1つの手段として、現在、トランスクリプショナルなプロファイルで何かそういうものが認められないかということを試んでいます。

従来の古典的なKnudsonのtwo-hit theoryでは、ここに癌抑制遺伝子 (tumor suppressor gene) というものがあつたとした場合に、そこに1つのミューテーションがあつて、その反対側の母由来・父由来という2つの染色体があつた場合に、片方が欠失を起こしている。そういうことで両方のコピーが不活化されることによって腫瘍が引き起こされるということですが、必ずしもこのように物理的に染色体が失われるわけではありません。ジェネティックにこの部分が欠けているということですが、こういうものは従来、マイクロサテライトの解析、CGH (Comparative Genomic Hybridization) という染色体コピー数を数えるような方法で、解析が可能でした。一方、エピジェネティック (epigenetic) な変化で、遺伝子のある領域がサイレンシングされている、ある特定の遺伝子が不活化されているという場合には、このLOHはマイナスで、コピー数としては2nが維持されているままですから、LOHとしては見つからないということが起こるわけです。

ここで遺伝子発現のインバランスをマップで見ようということで、1万、2万という遺伝子を解析したあとで、NCBIのGenbankのマップ情報にそれを貼り付けたわけです。

例えば肝癌ですが、染色体が1~22番、および性染色体が縦に並んでいます。先程の林崎先生の講演の中で、マウスのマップが1番から丸く並んでいた絵を思い出していただければ、似たようなことをここでやっているわけです。そして、ここの部分は肝細胞癌で落ちているような領域、こちらでは肝細胞癌で遺伝子が逆に上がっているという領域がマップしてあります。

そして、例えばここに白い線がありますのでここを

拡大してみると、短腕(p)で上がっているような部分、あるいは長腕(q)の方で上がっている部分、逆に下がっている部分が認められます。

先程、肝癌の分化度の話を少ししましたが、高分化(well differentiated)、中分化(moderately differentiated)の比較的正常に近い肝癌と、あるいはそこからさらに未分化(poorly differentiated)になった状態で比べます。つまり、高分化型から中分化型、中分化型から未分化型へ段階的に癌が悪性化していくわけですが、その間でこういう発現レベルの変化が染色体の領域として変わる部分はないかということです。

最初のこの絵は、この2つのことここをダイレクトに比べて、全体を見ているわけです。そうしますと、この領域は最初の段階ですでに発現亢進を行っていますが、この段階では変わりません。逆にこの部分は徐々に発現低下がaccumulationしていくことが理解できます。

この赤く示しているところに何があるかを見てみると、このような遺伝子が並んでいまして、この水色で示している部分、この星印がついているようなものが地図の上で確認できたものです。それ以外についても大体、肝細胞が通常発現しているような遺伝子が並んでいるだけで、現在のところは腫瘍にとりわけ関係があるというものは見当たりませんが、いわゆる癌化に伴って分化度が下がるに従って、発現が減っていきます。hepatocyte(肝細胞)の機能を担う遺伝子群がこの領域にクラスターしていた、同じファミリーに所属する遺伝子がこの領域にあったことを反映しているのだと思います。

こういう発現レベルをゲノムワイドに領域的に見ていくということで、もちろん染色体としての異常を検出することも期待できますし、あるいは今お示したような遺伝子のクラスターを認めることもできると思います。そして、そのステージによってどの段階で遺伝子発現の変化が起きていくかについても、見出すことができるのではないかと期待しています。

次に、こういう解析を診断に使えないかということで、肝腫瘍を例にとり、その手法について説明します。いろいろな分化度によって区別ができるであろう、あるいはC型肝炎、B型肝炎によって違うであろうということでしたが、それを識別するための遺伝子、実際にどういう遺伝子が違っているのかをどのようにして拾い出すかということで、ネイバーフッド(neighborhood)解析、あるいはnon-parametricなマン・ホイットニー(Mann-Whitney)解析を説明します。

ネイバーフッド解析というのは、ホワイトヘッド研究所のエリック・ランダーらのグループによって、2

年前に出た論文で提唱されたものです。このように2つの群、例えば正常と癌を比べた場合に、変化があるものを拾い出せばいいわけですから、なるべく平均値の差が大きくてばらつきが少ないものが、いいマーカーになるでしょう。その値をP値として定義して、そのP値の大きいものを選んできます。

使った症例はこのようなグループで、年齢以外は特に両群間で差はありません。肝炎の臨床をされている先生方にはご理解いただけるとと思いますが、B型肝炎とC型肝炎の患者さんというのは、年齢が高くなってから発症しますので、いたしかたがないところです。

またクラスター解析ですが、この非癌部組織のみをここに示しています。それもウイルス感染を伴った非癌部組織だけです。緑色がB型肝炎で、赤がC型肝炎です。先程の多変量解析の主因子分析の結果とほぼ同じで、B型肝炎とC型肝炎の発現パターンはそれぞれクラスターを作っているのがわかります。

この2つの間で、先程のP値の高い遺伝子を選びだそうということになります。実際に計算したP値が、全くランダムにそういうものが起こった場合に、どの程度のP値をとるだろうかということを計算した、計算上の全くランダムな場合と、さらに信頼区間を考慮した場合に、自分たちが計測した分布がどの程度有意かどうかということを検定します。それはpermutation testといいますが、このテストでこのように実際の測定した値が計算上のランダムな分布とP値の分布と変わっていなければ、それは有意ではないということになります。

この下のパネルを見ていただきますと、肝癌と非癌部組織のデータです。實際上、B型肝炎とC型肝炎ではランダムな状態よりも外れていますから、C型肝炎特有な遺伝子発現のパターンがあるだろうということかわかります。一方、こちら側の2つは癌組織HCCの中での分布ですが、両群間で差がないことがわかります。

こちら側がC型肝炎の非癌部組織で発現が亢進している遺伝子、こちら側がB型肝炎で発現が亢進している遺伝子です。この赤で示しているのはすべてインターフェロンで発現が亢進する遺伝子群です。つまり、C型肝炎ウイルスはRNAウイルスであり、それに対してB型肝炎ウイルスはDNAウイルスですから、C型肝炎の患者の肝組織ではRNAがウイルスに感染することによって起こる生体のレスポンスが、発現のプロファイルに反映していることがわかります。

一方、癌になった組織においては、そういうウイルスに対する応答は関係ないということで、あまり顕著な差がありませんでした。癌というのは継続的に変化

が蓄積してきますから、高分化型、中分化型、さらに低分化型の肝臓の中でどういう遺伝子の発現の変化が起きているかを見てみました。

先程のpermutation testを使って、高分化と中分化が区別できるかです。これは病理学的にも難しいことがありますが、かろうじて有意差がありそうです。そして、中分化と低分化については非常に差が大きそうでした。

遺伝子を拾い出してみますとこのようになり、この遺伝子のセットは高分化と中分化を比べて差のあるもの、つまり高分化で高く中分化で低いというもの、高分化で低く中分化で高いというものです。当然、中分化で高くなっているものは低分化でもそのまま高いですし、中分化で差があるものは、そのまま低分化でも差があることがここでもわかります。

こちらで抽出された約25個の約50個の遺伝子については、中分化と低分化の間でその差を見ているのですが、中分化から低分化へとより分化度が落ち、脱分化することによって、発現が下がるものがきれいに拾われてきています。こちらの25個の遺伝子についても、中分化から低分化になることによって上がっていきます。大体、こちらで下がっているものは、より分化度の高いものでも下がっていることがわかります。

このように見ていきますと、この約100個の遺伝子のパターンを見れば、分子レベルでも肝臓の分化度が大体判定できることがわかります。実際にこういう100個の遺伝子でスコアリングシステムを使えば、分化度についてはほぼ決定することができます。

最後にこういう情報をどのように実用化していくか。最近ではtranslational medicineなど盛んに言われますが、この中から新規の腫瘍マーカーが探索できないかということが行われます。

肝臓の場合にはAFP (α -フェトプロテイン) という、すでに臨床で長年使われているマーカーがあり、こういうものがDNAチップでどのように見えてくるかということになります。非癌部と肝細胞癌の組織を比べると、このように非常に突出しており、この線が30倍以上発現レベルが違う遺伝子のラインですから、既存の腫瘍マーカーであるAFPは、このようなRNAの発現レベルでも十分高いので拾われてきます。このように点は小さいですが、いくつか同様な動きを示しているものもあるので、そういうものについて、今、拾い出しを行ってみました。

AFPの問題点は、肝臓の患者だけではなく、その他の腫瘍でも数百程度までは上がる場合がありますし、肝硬変の状態や劇症肝炎からの回復期でもしばしば上昇が見られるということです。また、肝硬変の患者の場合には、特に小さい癌があるものではないかとい

うことで臨床の先生方の頭を悩ませています。そこでAFP以外のマーカー、現在はPIVKA IIなどが使われることがありますし、AFPの亜分画も使われています。その際に、よりたくさんのマーカーがあれば、より確実な早期診断が可能になります。

これは胃癌ですが、肝臓においてもいくつかの遺伝子の発現レベルが高分化、中分化、低分化で高くなっており、非癌部でも肝硬変で若干上がってくる傾向がありますが、このような遺伝子がいくつも見つかってきました。

先程のAFPをこういうマップで見ますと、高分化でほとんど上がっていません。臨床でご覧になられた場合にも、高分化の小さい肝臓では血清中のAFPレベルは上がってきませんが、この組織のRNAレベルを見てもAFPはあまり上がっていません。実際にこのチップの上でも、10例調べてせいぜい1例か2例しか上がっていませんでした。ですから、こういうものはより鋭敏なマーカーになる可能性があることが期待されます。

今の中の1つのモノクローナル抗体を作成して患者の腫瘍を見たところ、腫瘍でタンパク質がきちんと増えているのが確認されました。これはノーザンブロット、これはジーンチップのスコアですが、約50例の肝臓組織を検討して、6割ぐらいの患者でこのモノクローナル抗体についてはタンパク量の増加が認められています。

こちらの上の2例が陽性例で、こちらはネガティブですが、免疫組織染色をすると網の目状に見えているかがわかると思います。この網の目状は膜にタンパクがアソシエートしているということですが、実際上はこれは細胞外にあるタンパクですから、これが免疫療法ないしモノクローナル抗体によるミサイル療法に使えないかということも、現在、検討を進めています。

アフィメトリックスのシステムはこういうものです。ある程度信頼性がありますが、まだ臨床診断に使うには高価なシステムですし、自分の好きな遺伝子セットを自由に解析することはまだできませんので、フレキシビリティはありません。将来これが臨床検査で使われていくためには、もっと遺伝子の発現が簡単に測れるように、より高感度にしたものが望まれます。また、研究開発をするうえでも、さらに簡易なシステム、自分の好きな遺伝子をすぐにチップにできるシステムが望まれていくわけです。

これからこちらのセンターでも解析を行われるということですが、我々が感じている技術的な課題としては実験精度、再現性、感度で、まだ細胞に1遺伝子1コピーしかないというものが確実に測れるようなシステムではありません。また、臨床検体を測るという場合には、患者の貴重なサンプルでもありますから

微量検体です。しかも、腫瘍であってもその中で腫瘍成分、血球成分、血管の成分などがありますので、検体の均一性が問題になります。つまり、実験の細胞株でないかぎり、こういう組織の中のある部分の腫瘍とこちらの腫瘍では、また組織系が違うということもありますので、顕微鏡下にマイクロダイセクションすることも必要になってきます。

また、冒頭にマイクロアレイとジーンチップのシステムの両方をご紹介しましたが、その2つのプラットフォームの間でのデータの互換性もなかなか困難です。現在もその間での共通のデータベースが提唱されてはいますが、なかなか実現しないというのが世界的な現状です。加えて、技術的な問題としてもどのように標識をしていくか、RNAの増幅を行うか行わないかということや、その標識についても間接標識を行うかというさまざまな問題があります。これはどうにか使えるシステムにはなっていますが、今後、改良の余地がまだまだあるというのが実感です。

最後に機能情報、これはシステム情報学へということですが、今後、癌というものだけでも、さまざまなコンピューテーションなシミュレーションが期待されていきます。今日はDNAチップということを中心にお話ししましたが、プロテオームの解析情報、タンパク質の相互作用の情報などがデータベース化されれば、そういうものを理論的に肉付けをして、癌をシステムとして扱うことがいよいよ可能になってくるのではないかと期待しています。

例えばこれは β カテニンのパスウェイですが、先程出てきたAPC遺伝子や β カテニンといったものの発現レベルをパスウェイごとに一望することができれば、この癌で何が起きているかを理解しやすいインターフェースを作ることができます。

正常肝、肝癌をこのように並べてみますと、ここでは発現量を赤から黒へ g_{radual} に表示していても、人間の目にはよくわかりません。しかし、差分をとることによって、このように差が出てきて、発現量の低いものが青で、赤いところでは、例えばTCF-4という β カテニンからのシグナルが伝わる、直接、遺伝子のDNAに付く転写因子の発現が増えていることがわかります。

しかし、このパスウェイだけを見ても癌が理解できるわけではなく、例えばTGF β という抑制シグナルがsmadに入ってきて、このsmadがTCF-4のダウンストリームにある遺伝子の転写の調節を行います。こういうほかからのシグナル、あるいはこのシグナル系から外へのシグナルという、いわゆるシグナルのクロストークを考慮しないと生物は理解できません。こ

ういうインタラクションを、コンピュータの助けを借りながら理解していくことが非常に大事になると思いますし、特にこういう網羅的な情報を獲得して、その情報を整頓していくことがますます重要になってくると思います。

今日はゲノム情報とのカップリングをご説明申し上げましたし、今後は文献情報とのカップリングということで、生物情報のデータベースを作っていく。林崎先生からご紹介がありましたマウスのエンサイクロペディアが、まさにこういうものの代表になるわけですが、こういうデータは単なるデータです。今はデータの中からどうにか情報を抽出することまではできていますが、これにおぼれることなく、将来的には何らかの生物学的な発見をしたいと考えています。

これがうちのスタッフです。特に情報解析についてはこの広瀬先生の研究室、あるいは同じキャンパスにある岩田先生の研究室との共同で行っています。

座長：ありがとうございました。DNAチップを用いて癌を分類したり、診断するということですが、先生が指摘されたようにまだ問題点がかなりあるようです。しかし、おそらく近い将来には臨床で使われるのではないかと思います。質問等がありましたらお願いします。

参加者：微量サンプルの問題に少し興味がありますが、例えば臨床的にニードルバイオプシーでとれるようなものを扱って、メッセージを例えばランダムプライマーで増幅した解析結果と、リッチなメッセージをダイレクトに使った解析結果を比較した場合に、フォールスポジティブはどの程度出てきますか。また、どれくらいのサンプルを使えば信頼しうるデータになるのでしょうか。

油谷：RNAのレベルとして私たちがルーチンに使っているのは、トータルRNAで5マイクログラムです。細胞数にするとLCMでマイクロダイセクションで1万個が大体100ナノグラムですから、細胞として1マイクログラムが10の5乗ぐらいで、針の太さにもよりますが、ニードルバイオプシーでも十分可能な量にはなると思います。その場合に、臨床診断に使う部分などをいろいろ考えていきますと、なかなか倫理的に難しいところもありますが、もう少し量を減らしていけば現在のテクノロジーでも十分可能です。

ただし、減らしていった場合、特に現在、ランダムプライマーの場合には、アフィメトリックスのシステムは遺伝子の3'側に特にプローブの配列を用いています。cDNAのアレイと違い、ランダムプライマーで増えてきた配列の場合には、それがもともと使ってい

る配列よりももっと上流の方、5'側にバイアスされてしまう可能性があり、きちんとしたシグナルがとれない遺伝子が出てくる可能性はあると思います。それがこの場合には欠点になります。

ですから、アフィメトリックスのシステムの場合には、現在はT7RNAポリメラーゼを使って3'側から増幅をしています。それをさらに微量検体から増幅する場合には1回ランダムプライマーで戻します。そういう場合に若干バイアスが出ることがありますし、ランダムプライマーの反応でうまく反対側から戻ってこない場合に、約2割ぐらいの遺伝子は、発現が見られなくなる危険性は伴います。

参加者：癌の組織を採ってくる時、当然、正常組織の混入もあるから、レーザーキャプチャーのデータはどうされていますか。

油谷：肝癌の場合には、腫瘍の6~7割は腫瘍細胞なので。

参加者：やはり、「いいや」という感じですか。

油谷：ええ。言い訳になりますが、癌組織のプロファイリングということで許していただいて、癌組織を顕微鏡で見る場合には、癌細胞を取り囲む組織のリアクションも診断に入ってきます。実際に例えば内皮細胞で発現が亢進するというものもありますし、腫瘍血管特異的なトランスクリプトもありますので、今回のデータについてはLCMのデータではありません。

参加者：オリゴヌクレオチドのアフィ（アフィメトリックス）のチップのときに、ああいうアフィチップとオリゴヌクレオチドをプリンターで貼り付けたものとを比べたことはありますか。

油谷：あります。

参加者：どうですか。

油谷：オリゴヌクレオチドのどこを選ぶかによって、すぐくデータが変わります。傾向として増えているものは、アフィの結果と相関します。アフィも大体1年に1回ぐらいはプローブを見直して、見直すごとにだめなプローブをはじくので、徐々によくなっています。

参加者：少し違って、アフィと同じシーケンスをプリントで。テクニカルなことを聞いているのです。

油谷：アフィは発表すると言っていて、まだ25ベースのシーケンスを発表していないのです。（注：現在は公表している）

参加者：そうですか。遺伝子はわかっているのですね。

油谷：もちろん、遺伝子はわかっています。その25ベースの配列がどの領域か。

参加者：どこかがわからないのですか。

油谷：その数百ベースの中のどこかです、ということまでしかわかりません。ですから、cDNAのこの500ベースなら500ベースから、彼らは25塩基を十何か所選びましたと。その十何か所がことごとこということは、まだブラックボックスです。

参加者：それは言わないと、やはり・・・。

油谷：そういうプレッシャーが今は国際的に非常強くなって、去年から出すと言っていますが、いよいよ来年の早いうちに出すようです。そうすれば、またデータの検定が楽になりますし、自分たちでもう一度プローブのデータを計算し直すことが可能になります。

参加者：シーケンスがわからないらしいのですが、そういうデータは世の中にあるのですか。そこを知りたいのです。データというのはアフィのようなやり方でつないでいったチップと、別のところで合成機で作ったものを貼り付けていった同じシーケンスを見た場合に、データの質の差はあるでしょうか。

油谷：in situで合成するというのは、アジレントはインクジェットですが、そこで塩基を配って25塩基を合成していますから、あそこのデータはそれに近いのだと思います。それなりにロゼッタはデータを出していますし、それほど問題はないのではないかと思います。フォトリスグラフィという方法では、1サイクルごとの合成の効率が普通の合成機よりも悪いですから、実際にどれだけDNAがそこに合成されているかというのは、わからないのです。

そこが非常にミステリーで、林崎先生の質問にはきちんと答えられませんが、ほかのSAGE法やそのあとのRT-PCR、ノーザンで、私たちももともとは比べましたが、発現量の比較的多い遺伝子については、かなり絶対量を反映した相関関係が出ます。実際上のオリゴを貼り付けた場合についてのきちんとしたデータは配列がわかっていないので、アフィメトリックスは知っているかもしれませんが、ほかのところでは特にオープンになっているものはないと思います。

座長：ほかに質問はありませんか。なければ終わりたいと思います。油谷先生、ありがとうございました。